

PAPER • OPEN ACCESS

## Image defect detection algorithm based on deep learning

To cite this article: R A Sizyakin *et al* 2019 *IOP Conf. Ser.: Mater. Sci. Eng.* **680** 012041

View the [article online](#) for updates and enhancements.

## Image defect detection algorithm based on deep learning

R A Sizyakin<sup>1,3</sup>, V V Voronin<sup>1,2</sup>, N V Gapon<sup>1</sup>, A A Zelensky<sup>2</sup>, A Pižurica<sup>3</sup>

<sup>1</sup>Don State Technical University, sq. Gagarin 1 Rostov-on-Don, 344000 Russia

<sup>2</sup>Moscow State University of Technology “STANKIN”, ln. Vadkovsky 1 Moscow, 127055 Russia

<sup>3</sup>Ghent University, st. Rozier 44 Ghent Belgium,

E-mail: roma\_sz@mail.ru

**Abstract.** In this paper proposed a system for automatic defects detection in images. The solution to this problem is widely used in practice. Automatic detection is found in the challenge of detecting defects on the road surface, in the textile industry, as well as virtual restoration of archival photo images. The solution to this range of problems allows speeding up work in these areas, and in some cases, completely solving. To solve the first two problems (search for defects on the pavement and textiles), it is enough to create a mask that localizes defects in the image with maximum reliability, while photo restoration requires additional algorithms to restore the detected damaged areas. The proposed method is based on the latest achievements in the field of machine learning and allows solve the main disadvantages of traditional methods. Automatic defect detection is performed using a neural network with compound descriptor. A series of experiments confirmed the high efficiency of the proposed method in comparison with traditional methods for detecting defects.

### 1. Introduction

Solution automatic defect detection problem is widely used in practice. This problem occurs when searching for defects in the road surface, the textile industry, as well as virtual restoration of archival photo images. The decision of the range of tasks to speed up work in the direction of the data, and in some cases completely resolve. For defects occurring in the pictures are different kinds of stains, scratches, cracks and other foreign objects. Their appearance can be due to aging, physical impact, improper storage, or operation. Recovery of such defects in the present, in most cases carried out through manual processing, making it challenging to restore archived facsimiles.

In [1] for the detection of defects in the image using a method based on machine learning. For a mask with an estimated localization defects using morphological operations top and bottom hats to detect light and dark cracks. To reduce false marked areas on the pre-mask errors, using a method based on training. As the descriptor values are the hue, saturation, and brightness (HSV- Hue, Saturation, Value (Brightness)). For the construction of the separating hyperplane using a support vector machine (SVM) with a median radial basis function as the kernel (MRBF).



In [2] the same approach is used to detect defects that in [1]. The main distinguish this work from work [1] is the method of filling in damaged areas. Reconstruction of the damaged regions in this study occurs through search blocks similar to those damaged. As a similarity metric used the sum of squared differences (SSD). If the same unit was not found, then the bad block is filled with the average values. In [3] authors for detect defections proposed to use a morphological operation top hat, followed by binarization with the automatic selection of the threshold value [4]. Choice of the threshold value depends on the separability of the original image histogram. For optimal separation of the histogram iteratively selects threshold values that minimize the variance within a class, which is defined as a weighted sum of the variances of the two classes. In his work Otsu shown that minimizing the variance within a class is equivalent to maximize the variance between classes.

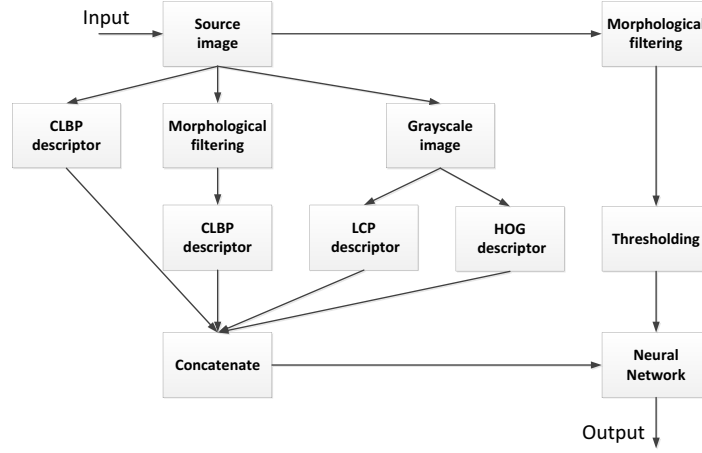
In [5] algorithm to search for defects in the photographs consists of three steps. The first is the improvement in pre-input image contrast. Then it calculates the convolution of the modified image from different directions Gaussian kernels. The second step is to obtain mask defects by applying morphological operations top hats with various sizes of structure-forming element. The third step is constructed for a pre-defect mask method using K-SVD. The point of this step is to train the algorithm, on pre-prepared templates, followed by the classification of areas on the original image. To reduce the number of false positives in this paper, the method to-medium. The handle acts as a color component, the length, orientation and eccentricity ratio. After that three-prepared masks are combined into one by adding. In those areas where there is the imposition of the alleged defects, the resultant mask takes the single value, or zero.

In this paper, we propose a new approach to detection defects on the photograph. The defect detection method includes two main stages: preliminary localization of defects and subsequent accurate classification of detected defects.

## 2. Proposed method

Damages encountered in the photo image include various kinds of spots, scratches, cracks and other foreign objects. Their appearance can be caused by aging, physical stress, improper storage or use. In this paper, we focus on crack detection. The method for detecting damage to an image consists of two main steps: the first step is to use morphological filtering as a pre-processing, the second step is to use the machine learning method, which is necessary to classify pixels that have received a large response in the preprocessing phase (see Fig.1).

Morphological filtering reduces the computational complexity of the algorithm, because there is no need to process each pixel of the image, but only those that received a higher response. Also, the use of this operation can reduce the number of false alarms [6].



**Figure 1** General scheme of the proposed method of defect detection.

In our work we perform morphological operation "bottom hat". The essence of this operation consists of subtraction from the original image  $I_{r,c}$  the result of its "opening" with the structural element  $B$ . Use of this morphological operation allows locating the expected defects in the image.

$$Mask_{r,c} = Y_{r,c} - (Open_{r,c})^B \quad (1)$$

$$(Open_{r,c})^B = (Y_{r,c} \ominus B) \oplus B \quad (2)$$

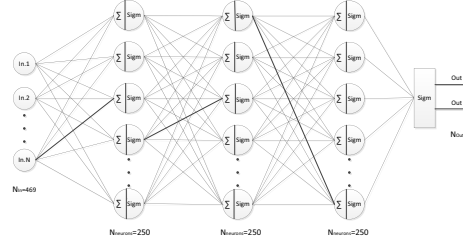
$$(Dilate_{r,c})^B = MAX_{(u,v) \in B} (Y_{r,c}(r+u, c+v) - B(u, v)) \quad (3)$$

$$(Erode_{r,c})^B = MIN_{(u,v) \in B} (Y_{r,c}(r+u, c+v) - B(u, v)) \quad (4)$$

where  $B$  is a structural element with size pixel size  $u \times v$ ,  $\ominus$  - erode,  $\oplus$  - dilate.

After the procedure of morphological filtering, the preliminary mask with defects contains both correctly detected defects and a number of false positives. To reduce the number of false alarms, we use a neural network as a classifier.

The neural network allows reducing false alarms on the preliminary mask obtained by morphological filtering. The neural network architecture for defect detection is shown in figure 2.



**Figure 2.** Neural network architecture for damage detection

The network consists of 3 hidden layers. Each layer includes 250 neurons. The logistic sigmoid is used as the activation function:

$$f(x) = \frac{2}{1 + e^{-x}} - 1 \quad (5)$$

where  $x$  - feature vector.

To determine the losses, we use the binary cross-entropy function. This is defined according to the expression:

$$H(y, y') = -\frac{1}{m} \sum_{l=1}^m [y \cdot \log(y') + (1-y) \cdot \log(1-y')] \quad (6)$$

where  $y$  - prediction Conv.Net,  $y'_j$  - true value.

Additional we use the method of optimization Adam, proposed in work [7], with learning rate equal 0.0005. Training it took approximately 45 epochs. We concatenated to one and used the following descriptors as input for neural network: CLBP [8], HOG [9], and LCP [10].

Texture CLBP operator is an extension of work [11]. The main difference lies in the preservation of not only the sign, but also the magnitude of the component:

$$d_p = s_p * m_p, \begin{cases} s_p = \text{sign}(d_p) \\ m_p = |d_p| \end{cases}, \quad s_p = \begin{cases} 1 & d_p \geq 0 \\ -1 & d_p < 0 \end{cases} \quad (7)$$

where  $d_p$  is the vector of difference between the Central pixel and its neighbors,  $s_p$  is the vector of the sign component,  $m_p$  is the vector of the magnetic component.

The use of two components allows achieving the accuracy of the texture description higher than the original LBP method, which uses only the sign component. Texture operator LCP is a combination of feature vectors of the LBP method and weighted coefficients (MiC). Coefficients are calculated for adjacent pixels to the center, in the LBP pattern. The formula for calculating the coefficients is written as follows:

$$E(a_0, \dots, a_{P-1}) = \left| g_c - \sum_{r=0}^{P-1} a_r g_r \right| \quad (8)$$

where  $g_c$  is the center pixel,  $g_r$  is the adjacent pixel,  $a_r$  is the weight coefficient calculated for each  $g_r$ .

To calculate the weights  $a_r$ , the least-squares method is used:

$$C_L = V_L A_L \quad (9)$$

$$A_L = (V_L^T V_L)^{-1} L_L^T C_L \quad (10)$$

where  $C_L$  is the vector of the types of patterns of interest,  $V_L$  is the intensity (brightness) of adjacent pixels,  $A_L$  is the vector with unknown values of the weight coefficients  $a_r$ ,  $L$  is the type of pattern.

Image rotation stability is achieved by applying the Fourier transform to a vector  $A_L$ :

$$H_L(k) = \sum_{i=0}^{P-1} A_L(i) \cdot e^{-j2\pi k i / P} \quad (11)$$

where  $A_L(i)$  is  $i$  the vector element  $A_L$ .

Leaving the amplitude component of the vector, we get an additional vector (MiC) for the original LBP method. Texture operator HOG based on counting the number of gradient directions in local areas of the image, on a dense grid of uniformly distributed cells. To increase the accuracy of the description of the texture features of the image, normalization of local contrast is used. The algorithm for constructing the feature vector consists of several stages. At the first stage, the values of the gradients are calculated. To calculate the gradients, a one-dimensional differentiating mask is used in the horizontal and vertical directions. At the next stage, the calculated gradients are used to form a histogram for each of the cells, passing the weighted voting procedure.

In order to take into account, the brightness characteristics of the image, the gradients are locally normalized. For this, cells are grouped into larger connected blocks. Normalization of blocks occurs in accordance with the expression:

$$f = \frac{v}{\sqrt{\|v\|_2^2 + e^2}} \quad (12)$$

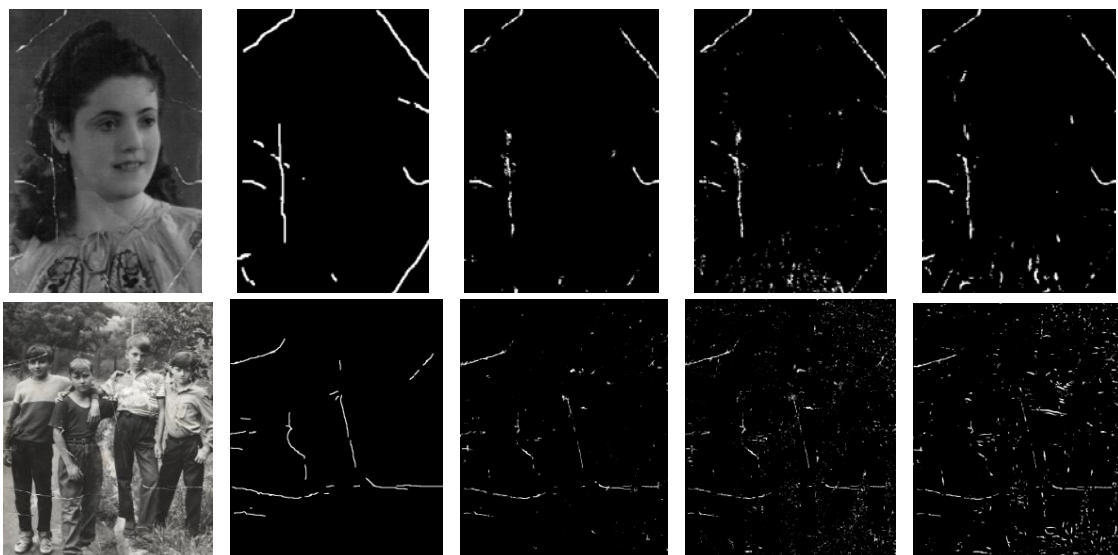
where  $f$  is the normalization factor,  $v$  is the non-normalized vector,  $\|v\|_2$  is its norm,  $e$  is a certain small constant.

As a result, the resulting descriptor has a length of 469 bins, of which 108 values correspond to the values of the CLBP operator, which is calculated for each color component of the color image, 36 and 289 values, correspond to the texture operators HOG and LCP, respectively, calculated for the original image in grayscale, and also 36 values correspond to the values of the CLBP operator calculated for the image after morphological filtering. This composite descriptor is calculated for a local window with  $20 \times 20$  the pixel size taken around each pixel marked by a unit after threshold processing.

### 3. Results and discussion

To compare the effectiveness of detecting damage in the image, the proposed method is compared with methods based on machine learning: convolutional neural networks (CNN) and support vector machine method (SVM). Three color components of the image are used as input to the convolutional neural network. The first convolutional layer has 10 feature maps, the second layer has 20 feature maps, the third layer has 30 feature maps. The size of the convolution filters is  $5 \times 5$  pixels. ReLU is used as an activation function in hidden layers. The following parameters were also used: the learning rate is 0.001, the size of the training data batch is 20 samples, for the training was using the ADAM method. The support vector machine used a linear separating hyperplane, with an acceptable error of 5%. The previously described texture descriptors are used as a descriptor.

Figure 3 illustrates the result of the proposed method for detecting defects, a method based on support vector machine and the method based on the use of convolutional neural networks.



**Figure 3.** The first column corresponds to images with defects; the second column – ground truth; the third column – proposed method; the fourth column – SVM; the fifth column – CNN.

Approximately 2,500 samples ( $20 \times 20$  pixels) containing cracks as well as 2,500 undamaged samples ( $20 \times 20$  pixels) were used for training all methods. Table 1 shows the results for 5 test images. Test images obtained from free access.

**Table 1.** The results of determining the iron content in the wear particles.

	Probability false alarm			Probability correct detection
	Proposed method	CNN	SVM	
Test image 1	0.035	0.056	0.053	0.92
Test image 2	0.027	0.073	0.063	0.83
Test image 3	0.051	0.053	0.115	0.79
Test image 4	0.00215	0.00247	0.00237	0.6
Test image 5	0.0074	0.0110	0.0138	0.5
Averaged	0.0251	0.0398	0.0432	0.65

From the estimates given in table 1, it can be concluded that the proposed method reduces the number of false positives by an average of 1.6 times, while maintaining the probability of correct detections on a par with known techniques. The disadvantages of the classifier based on the convolutional neural network include the fact that it has a low generalizing ability with not enough training data. Also, an excessive increase in the training parameters can further reduce the generalization ability. The main drawback of the support vector machine classifier is that it is not efficient enough to classify using a multi-part descriptor that includes multiple texture operators.

#### 4. Conclusion

In this work it was proposed method for the automatic detection of defects on the image. For a preliminary localization of the crack on the image in this work, we used morphological filtering. To reduce the number of false alarms is used a neural network with concatenating vector descriptors. of the proposed method with the various known methods, showed high efficiency in the use of neural networks for detection of defects in the image. The result probably can be improved depending on particular tasks, for example by increasing the size of the training data, modifying the architecture neural networks, changing set descriptors or set parameter more accurate, the resolution of the data used can also be increased, increasing contrast, etc.

#### 5. Acknowledgments

This work was supported by Russian Ministry of Education and Science in accordance to the Government Decree № 218 from April 9, 2010 (project number № 074-11-2018-013 from May 31, 2018 (03.G25.31.0284)).

#### References

- [1] I Giakoumis, N Nikolaidis, I Pitas. Digital image processing techniques for the detection and removal of cracks in digitized paintings. *Department of Informatics Aristotle University of Thessaloniki*
- [2] S G Schirripa, F Somma. 2010 Virtual restoration of cracks in digitized image of paintings. *International Conference on Defects in Insulating Materials, Journal of Physics: Conference Series* **249** 012059
- [3] A Gupta, V Khandelwal, A Gupta, M C S Thammastat 2008 Image Processing Methods for the Restoration of Digitized Paintings. *Int. J. Sc. Tech.* **Vol.** 13, No.3

- [4] N Otsu 1979 A Threshold Selection Method from Gray-Level Histogram. *IEEE Transaction on Systems, Man, and Cybernetics*, **vol.** SMC-9, no.1, pp. 62-66
- [5] B Cornelis, T Ruzic, E Gezels, A Doods, A Pizurica, L Platasa, J Cornelis, M Martens, M De Mey, I Daubechies. 2013 Crack detection and inpainting for virtual restoration of paintings: The case of the Ghent Altarpiece. *Signal process.*, **vol.** 93, no3, pp. 605-619
- [6] V Voronin, V Marchuk, R Sizyakin, N Gapon, M Pismenskova, S Tokareva. 2016 Automatic image cracks detection and removal on mobile devices. *Mobile Multimedia/Image Processing, Security, and Applications*
- [7] Kingma D and Ba J Adam 2014 A method for stochastic optimization. *CoRR*
- [8] Z Guo, L Zhang and D Zhang 2010 A completed modeling of local binary pattern operator for texture classification. *IEEE Transactions on Image Processing*, **vol.** 19, no. 6, pp. 1657–1663
- [9] N Dalal and B Triggs. 2005 Histograms of oriented gradients for human detection. **INRIA**
- [10] Yimo Guo, Guoying Zhao, and Matti Pietikainen. 2011 Texture Classification using a Linear Configuration Model based Descriptor. *Proceedings of the British Machine Vision Conference*, pp. 119.1-119.10
- [11] Ojala T, Pietikainen M, Maenpaa T 2002 Multiresolution Grayscale and Rotation Invariant Texture Classification with Local Binary Patterns. *IEEE Transactions on pattern analysis and machine intelligence*, **vol.** 24, no. 7